# Part 1. Classification 分类

## Chapter 1. Machine learning basics 机器学习基础

1、机器学习：利用计算机来彰显数据背后的真正含义。

2、机器学习类型：

3、各类学习的用途：

4、开发机器学习应用程序步骤：

## Chapter 2. Classifying with k-Nearest Neighbors k-近邻算法

1、概述：（范围：数值型&标称型）（缩写kNN）

优点：精度高、对异常值不敏感、无数据输入假定。

缺点：计算复杂度高、空间复杂度高。

2、 ①计算已知类别数据集中的点与当前点之间的距离；

②按照距离递增次序排序；

③选取与当前点距离最小的k个点；

④确定前k个点所在类别的出现频率；

⑤反馈前k个点出现频率最高的类别作为当前点的预测分类。

## Chapter 3. Splitting datasets one feature at a time: decision trees

## Chapter 4. Classifying with probability theory: naïve Bayes

## Chapter 5. Logistic regression

## Chapter 6. Support vector machines

## Chapter 7. Improving classification with the AdaBoost meta-algorithm

# Part 2. Forecasting Numeric Values With Regression

## Chapter 8. Predicting numeric values: regression

## Chapter 9. Tree-based regression

# Part 3. Unsupervised Learning

## Chapter 10. Grouping unlabeled items using k-means clustering

## Chapter 11. Association analysis with the Apriori algorithm

## Chapter 12. Efficiently finding frequent itemsets with FP-growth

# Part 4. Additional Tools

略